

# Pembangkitan interpretasi tekstual berbahasa Indonesia berdasarkan data pemeriksaan kimia darah menggunakan pendekatan berbasis r-template

## *Textual interpretation generation in Indonesian language based on blood chemistry tests data using r-template based approach*

Indra Aulia<sup>\*)</sup>, Sarah Purnamawati, Junianto

Program Studi Teknologi Informasi, Universitas Sumatera Utara  
Jalan Alumni No. 3, Kampus USU Padang Bulan, Medan, Indonesia 20155

**Cara sitasi:** I. Aulia, S. Purnamawati, and J. Junianto, "Pembangkitan interpretasi tekstual berbahasa Indonesia berdasarkan data pemeriksaan kimia darah menggunakan pendekatan berbasis r-template", *Jurnal Teknologi dan Sistem Komputer*, vol. 8, no. 1, pp. 44-53, 2020. doi: [10.14710/jtsiskom.8.1.2020.44-53](https://doi.org/10.14710/jtsiskom.8.1.2020.44-53), [Online].

**Abstract** - The result of the blood chemistry tests is usually presented in the form of a table written in abbreviations, numbers, and units. Unfortunately, the young doctors often require the time and experience for interpreting the blood chemistry tests into a textual representation, which is easy to read and understand. Therefore, this research aimed at developing a system (BTISys) that can generate the textual representation in the Indonesian language automatically based on the blood chemistry test. BTISys generates the representation using Natural Language Generation (NLG) approach based on the r-template method. The reliability of BTISys is measured by considering the naturalness of generated textual representation. The naturalness can be observed by three criteria, such as readability, clarity, and general appropriateness. The reliability of BTISys is quite good to generate the textual representation automatically. It can be seen from the readability, clarity, and general appropriateness, which reach 73 %, 70 %, and 60 % respectively, that implies the naturalness of generated textual representation.

**Keywords** – blood chemistry tests; natural language generation; template based; naturalness; Indonesian language

**Abstrak** – Hasil pemeriksaan laboratorium kimia darah umumnya direpresentasikan dalam bentuk tabel yang ditulis dalam singkatan dan angka yang dilengkapi dengan satuannya. Dokter muda biasanya membutuhkan waktu dan pengalaman dalam menginterpretasikan hasil pemeriksaan tersebut ke dalam representasi tekstual yang mudah dibaca dan dipahami. Penelitian ini bertujuan untuk membangun suatu sistem (BTISys) yang mampu membangkitkan representasi tekstual dalam bahasa Indonesia secara otomatis berdasarkan hasil pemeriksaan laboratorium kimia darah. BTISys membangkitkan representasi

tersebut dengan menggunakan pendekatan Natural Language Generation (NLG) berbasis metode r-template. Keandalan BTISys pada penelitian ini diukur berdasarkan tingkat kealamian representasi tekstual yang dibangkitkan. Tingkat kealamian tersebut dapat dilihat melalui tiga kriteria, yaitu keterbacaan, kejelasan dan kesesuaian umum. Berdasarkan hasil pengujian dapat disimpulkan bahwa BTISys memiliki keandalan yang cukup baik dalam menghasilkan suatu interpretasi ke dalam representasi tekstual secara otomatis menggunakan bahasa Indonesia. Hal ini terlihat dari tingkat keterbacaan yang mencapai 73 %, kejelasan 70 %, dan kesesuaian umum 60 % yang menyiratkan kealamian dari representasi tekstual yang dibangkitkan oleh BTISys.

**Kata kunci** – pemeriksaan kimia darah; natural language generation; template based; naturalness; bahasa Indonesia

### I. PENDAHULUAN

Kesehatan manusia umumnya terbagi atas 2 (dua) unsur, yaitu unsur kesehatan psikis (rohani) dan unsur kesehatan fisik (jasmani). Unsur kesehatan fisik (jasmani) menjadi salah satu unsur yang diperhatikan dan dipertimbangkan oleh tim medis, seperti dokter, dalam rangka untuk mengetahui kondisi kesehatan manusia, baik ketika pemantauan kesehatan medis rutin (*medical checkup*) atau ketika terdapat indikasi penyakit tertentu. Kondisi unsur kesehatan fisik (jasmani) ini dapat diketahui melalui berbagai proses pemeriksaan, salah satunya adalah melalui pemeriksaan laboratorium.

Dalam dunia medis, terdapat berbagai jenis pemeriksaan laboratorium. Salah satu jenis pemeriksaan yang paling sering digunakan untuk menilai kondisi kesehatan umum, mengkonfirmasi adanya infeksi, dan melihat kondisi organ-organ tertentu adalah pemeriksaan darah atau hematologi. Pemeriksaan hematologi ini umumnya terbagi atas beberapa jenis pemeriksaan, di antaranya hitung darah lengkap (HDL),

<sup>\*)</sup> Penulis korespondensi (Indra Aulia)  
Email: [indraaulia@usu.ac.id](mailto:indraaulia@usu.ac.id)

kimia darah, enzim darah, dan pemeriksaan darah untuk menilai resiko penyakit jantung [1].

Pada penerapannya, hasil pemeriksaan kimia darah biasanya direpresentasikan dalam bentuk tabel. Representasi ini umumnya hanya berisi identitas pasien, data pemeriksaan dengan nama komponen, dan keterangan yang sebagian ditulis dalam singkatan dan angka-angka yang dilengkapi dengan satuannya, seperti glukosa darah 94 mg/dL, albumin 2,5 g/dL, dan *blood urea nitrogen* (BUN) 94 mg/dL. Namun, representasi ini tidak disertakan dengan penjelasan tekstual. Hal ini menyebabkan hasil pemeriksaan laboratorium kimia darah membutuhkan waktu yang relatif lama untuk diinterpretasikan oleh dokter muda atau tenaga medis. Hal ini juga dapat menimbulkan kesalahan interpretasi. Oleh karena itu, sistem yang dapat menginterpretasi hasil laboratorium kimia darah ke dalam representasi tekstual untuk membantu dokter muda atau tenaga medis dalam mempercepat proses interpretasi data yang biasanya dilakukan secara manual.

Penelitian yang berfokus pada pembangkitan representasi tekstual berdasarkan data non tekstual secara otomatis telah banyak dilakukan [2]-[8]. Penelitian tersebut secara umum menggunakan pendekatan *Natural Language Generation* (NLG) untuk berbagai domain masalah menggunakan berbagai pendekatan dan metode, di antaranya pembangkitan tekstual ramalan cuaca berbasis metode *Case Based Reasoning* (CBR) dilakukan dalam [2], kajian penerapan pengetahuan geografis dari web untuk membangkitkan deskripsi jaringan sensor hidrologi bagi publik [3], dan pembangkitan teks dari masukan non-tekstual (dataset tabular) [4].

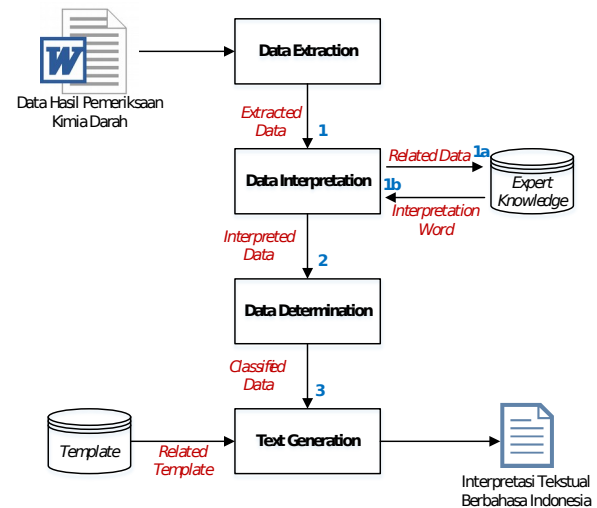
Secara khusus untuk bidang kesehatan, pembangkitan ringkasan tekstual dari data medis neonatal di ruang ICU menggunakan pendekatan *data-to-text* dilakukan dalam [5], pembangkitan ringkasan rawat inap pasien dalam [6], pembangkitan interpretasi tekstual berbahasa Indonesia berdasarkan data grafik pengawasan kesehatan satu garis menggunakan metode bigram dalam [7], dan pembangkitan interpretasi berbahasa Indonesia berdasarkan data grafik pengawasan kesehatan dua garis menggunakan metode template dalam [8]. Sejalan dengan penerapan di bidang kesehatan tersebut, penelitian ini bertujuan dan berfokus dalam mengembangkan sistem interpretasi tekstual berbahasa Indonesia berdasarkan data hasil pemeriksaan kimia darah secara otomatis dengan menggunakan pendekatan NLG berbasis metode *r-template*.

## II. METODE PENELITIAN

Dalam rangka untuk mengembangkan suatu sistem pembangkitan representasi tekstual bahasa Indonesia berdasarkan data hasil pemeriksaan kimia darah, penelitian ini mengusulkan suatu sistem yang disebut dengan *Blood Test Interpreter System* (BTISys). Sistem ini mengadopsi konsep dasar sistem NLG pada umumnya. Konsep rancangan arsitektur BTISys ditunjukkan pada Gambar 1.



Gambar 1. Konsep rancangan arsitektur BTISys



Gambar 2. Arsitektur umum BTISys yang diusulkan

Konsep rancangan arsitektur BTISys dinyatakan menjadi suatu arsitektur umum sistem yang lebih rinci (Gambar 2). BTISys terdiri atas beberapa proses, yaitu ekstraksi data, interpretasi data, penentuan data, dan pembangkitan teks. Masukan BTISys adalah suatu dokumen berformat Microsoft Word yang berisi data hasil pemeriksaan kimia darah, seperti data hasil pemeriksaan kadar gula darah, fungsi hati, fungsi ginjal, dan fungsi lemak. Data tersebut diproses melalui 4 (empat) proses yang ada di BTISys sehingga menghasilkan suatu keluaran berupa file *Portable Document Format* (PDF) yang berisi representasi tekstual berbahasa Indonesia sebagai wujud dari interpretasi data hasil pemeriksaan kimia darah. Keluaran ini disajikan bersamaan dengan data hasil pemeriksaan kimia darah. Keluaran tersebut disisipkan di bawah data hasil pemeriksaan kimia darah sebagai wujud penjelasan terhadap data terkait.

### A. Masukan

Masukan BTISys yang digunakan pada penelitian ini adalah data hasil pemeriksaan kimia darah yang diperoleh dari suatu alat *hematology analyzer*. Alat ini umumnya menampilkan hasil pemeriksaan kimia darah ke suatu layar monitor. Hasil pemeriksaan tersebut disajikan ke dalam bentuk struktur teratur, yaitu tabel, di suatu dokumen berformat Microsoft Word yang diperoleh dari salah satu rumah sakit. Contoh isi dokumen ditunjukkan pada Gambar 3.

Data hasil pemeriksaan kimia darah yang diinterpretasikan ke dalam suatu representasi tekstual berbahasa Indonesia meliputi data hasil pemeriksaan kadar gula darah, fungsi hati, fungsi ginjal, dan fungsi

Nama Pasien : Akbar Duwi  
Tanggal Lahir : 4 Maret 1979  
Jenis Kelamin : Pria

Dokter Pengirim : dr. Dina Mutia  
Tanggal Pemeriksaan : 26 Juli 2016  
Nomor Pemeriksaan : KD260716022

JENIS PEMERIKSAAN	HASIL PEMERIKSAAN	NILAI RUJUKAN
<b>KIMIA DARAH</b>		
<b>DIABETES</b>		
Glucose	86	70 - 180 mg/dl
<b>FUNGSI HATI</b>		
SGOT	24	W : ≤ 31 u/l P : ≤ 35 u/l
SGPT	26	W : ≤ 34 u/l P : ≤ 45 u/l
<b>FUNGSI GINJAL</b>		
Ureum	90	10 - 50 mg/dl
Creatinine	1.1	W : 0.5 - 1.1 mg/dl P : 0.6 - 1.4 mg/dl
Asam Urat	5.2	P : 3.5 - 7.2 mg/dl W : 2.9 - 5.2 mg/dl
<b>LEMAK</b>		
Cholesterol Total	290	< 200 mg/dl
HDL Cholesterol	39	W : 35 - 70 mg/dl P : 35 - 55 mg/dl
LDL Cholesterol	180	< 130 mg/dl
Trygliserida	200	< 150 mg/dl
<b>IMUNOLOGI</b>		
HBsAG	NON REAKTIF	Negatif

Catatan :

Pemeriksa,

dr. Muhammad Putra, Sp.PK

**Gambar 3.** Contoh masukan BTISys

lemak. Secara teknis, dokumen data hasil pemeriksaan kimia darah terbagi atas 3 (tiga) bagian utama, yaitu data diri pasien (seperti nama, tempat/tanggal lahir, umur, jenis kelamin, tanggal pemeriksaan dan dokter pengirim pasien), data hasil pemeriksaan, dan pengesahan petugas medis. Namun, bagian pertama dan bagian ketiga merupakan data yang disamakan dalam rangka menghormati dan menjaga kode etik medis yang berlaku terkait data. Sementara itu, bagian kedua merupakan data hasil pemeriksaan kimia darah yang disajikan dalam tabel terstruktur dan terbagi atas 4 (empat) komponen pemeriksaan kimia darah. Masing-masing komponen tersebut memiliki 2 (dua) kolom penyerta yang menampilkan data hasil pemeriksaan dari pasien tertentu dan nilai normal yang menjadi rujukan dasar untuk setiap komponen pemeriksaan.

## B. Ekstraksi data

Proses pertama dalam BTISys adalah ekstraksi data (*Data Extraction*). Proses ini bertanggung jawab untuk membaca dan mengekstrak data hasil pemeriksaan kimia darah yang terdapat dalam suatu dokumen berformat Microsoft Word seperti yang dijelaskan pada Bagian A. Dalam masalah ini, data yang dibaca dan diekstraksi adalah data diri pasien dan data hasil pemeriksaan kimia darah. Hasil dari proses ini adalah suatu *List* yang berisi kumpulan variabel dan nilai yang disimpan secara terstruktur (**Gambar 4**) sehingga mudah dimanipulasi pada proses berikutnya. Hasil ini dinyatakan sebagai *Extracted Data*.

Untuk membaca dan mengekstraksi data dari dokumen berekstensi Microsoft Word, penelitian ini

```
{
  {
    'Nama Pasien' : 'Akbar Duwi',
    'Jenis Kelamin' : 'Pria',
    'Pemeriksaan' : {
      {
        'Jenis Pemeriksaan' : 'Glucose',
        'Hasil Pemeriksaan' : 86
      },
      {
        'Jenis Pemeriksaan' : 'SGOT',
        'Hasil Pemeriksaan' : 24
      },
      {
        'Jenis Pemeriksaan' : 'SGPT',
        'Hasil Pemeriksaan' : 26
      },
      {
        'Jenis Pemeriksaan' : 'Ureum',
        'Hasil Pemeriksaan' : 90
      },
      {
        'Jenis Pemeriksaan' : 'Creatinine',
        'Hasil Pemeriksaan' : 1.1
      },
      {
        'Jenis Pemeriksaan' : 'Asam Urat',
        'Hasil Pemeriksaan' : 5.2
      },
      {
        'Jenis Pemeriksaan' : 'Cholesterol Total',
        'Hasil Pemeriksaan' : 290
      },
      {
        'Jenis Pemeriksaan' : 'HDL Cholesterol',
        'Hasil Pemeriksaan' : 39
      },
      {
        'Jenis Pemeriksaan' : 'LDL Cholesterol',
        'Hasil Pemeriksaan' : 180
      },
      {
        'Jenis Pemeriksaan' : 'Trygliserida',
        'Hasil Pemeriksaan' : 200
      },
      {
        'Jenis Pemeriksaan' : 'HBsAG',
        'Hasil Pemeriksaan' : 'Non Reaktif'
      }
    }
  }
}
```

**Gambar 4.** Keluaran proses ekstraksi data (*Extracted Data*)

## Algoritme 1. Pseudocode ekstraksi data

```
1 Input : Dokumen File Word berisi Tabel as FWord
2 Output: Extracted Data
3 Function DataExtraction (FWord)
4   ExtractedData = []
5   foreach Row ∈ TableInFWord
6     foreach Cell ∈ Row
7       if cell(key) = (nama or jenis
         kelamin) Then masukkan cell(key) dan
         cell(value) ke ExtractedData
8       else
9         masukkan cell(key) dan cell(value) ke
         ExtractedData sebagai anggota list
         key pemeriksaan.
10      end foreach
11    end foreach
12 end function
```

telah memformulasikan suatu metode *Data Extraction* seperti dinyatakan pada **Algoritme 1**. Metode ini menghasilkan suatu keluaran seperti contoh pada **Gambar 4**.

## C. Interpretasi data

*Extracted Data*, yang merupakan keluaran dari proses Data Extraction, selanjutnya diteruskan ke proses interpretasi data (*Data Interpretation*). Proses ini bertanggung jawab untuk menemukan dan menentukan interpretasi awal dan interpretasi lainnya dari setiap data komponen pemeriksaan kimia darah. Interpretasi awal mengacu pada interpretasi data pemeriksaan apakah

normal, meningkat, dan menurun. Interpretasi lainnya meliputi jenis kata interpretasi awal dan rentang nilai (jika data menunjukkan abnormal). Contoh keluaran hasil interpretasi data ditunjukkan dalam [Gambar 5](#).

Penentuan interpretasi tersebut dapat dibangkitkan oleh BTISys berdasarkan pengetahuan pakar dan berbagai sumber literatur yang telah terformulasikan dan tersimpan dalam suatu *knowledge base*. Pengetahuan pakar yang dimaksud mengacu kepada rentang nilai rujukan normal dan abnormal untuk setiap komponen pemeriksaan kimia darah. Rentang nilai normal komponen pemeriksaan kimia darah yang digunakan pada penelitian ini mengacu kepada beberapa standar sebagaimana yang ditunjukkan pada [Tabel 1](#).

Komponen pemeriksaan kimia darah secara umum dapat diinterpretasikan dalam 3 (tiga) jenis interpretasi, yakni normal, meningkat, dan menurun. Aturan interpretasi ini dapat dilihat pada [Tabel 2](#). Berdasarkan aturan tersebut, interpretasi data menentukan interpretasi awal untuk setiap komponen pemeriksaan kimia darah. Metode interpretasi data dapat diproses menggunakan metode yang disajikan pada [Algoritme 2](#). Perhitungan besaran selisih nilai hasil pemeriksaan dengan nilai rujukan yang telah ditetapkan sesuai standar dapat dihitung menggunakan Persamaan 1. Parameter  $DifVal$  menyatakan besaran nilai perubahan,  $NHP$  nilai hasil pemeriksaan,  $NR_{ub}$  nilai rujukan batas atas, dan  $NR_{lb}$  nilai rujukan batas bawah.

$$DifVal = \begin{cases} NHP - NR_{ub}, & NHP > NR_{ub} \\ NR_{lb} - NHP, & NHP < NR_{lb} \\ 0, & NR_{lb} \leq NHP \leq NR_{ub} \end{cases} \quad (1)$$

#### D. Penentuan data

*Interpreted Data* yang merupakan keluaran dari proses interpretasi data diteruskan ke proses selanjutnya, yaitu pemilihan data (*Data Determination*). *Data Determination* bertanggung jawab untuk memilih data mana yang memiliki urgensi penting untuk disampaikan nantinya dalam representasi tekstual. Kriteria pemilihan urgensi ini tentunya berdasarkan interpretasi awal yang telah dibangkitkan pada proses sebelumnya menggunakan pengetahuan terkait.

Kriteria tersebut adalah setiap data yang interpretasi awal yang dibangkitkan bukan merupakan suatu kondisi normal, maka data tersebutlah yang dipilih. Hal ini disebabkan karena dalam hasil pemeriksaan kimia darah, konten atau pesan yang perlu dan penting disampaikan adalah konten atau pesan dari data komponen pemeriksaan yang nilai pemeriksaannya berada di luar dari batas nilai normal. Oleh karena itu, keluaran dari proses *Data Determination* merupakan data yang perlu dan penting untuk disampaikan.

Keluaran tersebut disebut dengan *Selected Data* seperti yang disajikan pada [Gambar 6](#). Dengan kata lain, keluaran tersebut adalah data yang nilai pemeriksaannya berada di bawah atau di atas nilai normal untuk setiap komponen pemeriksaan kimia darah. Proses *Data*

```
{
  'Jenis Pemeriksaan' : 'Creatinine',
  'Hasil Pemeriksaan' : 1.1,
  'Interpretasi'      : 'Normal',
  'I_type'            : 'Noun',
  'DifVal'            : 0
},
{
  'Jenis Pemeriksaan' : 'Asam Urat',
  'Hasil Pemeriksaan' : 5.2,
  'Interpretasi'      : 'Normal',
  'I_type'            : 'Noun',
  'DifVal'            : 0
},
{
  'Jenis Pemeriksaan' : 'Cholesterol Total',
  'Hasil Pemeriksaan' : 290,
  'Interpretasi'      : 'Peningkatan',
  'I_type'            : 'Noun',
  'DifVal'            : 0
},
{
  'Jenis Pemeriksaan' : 'HDL Cholesterol',
  'Hasil Pemeriksaan' : 39,
  'Interpretasi'      : 'Normal',
  'I_type'            : 'Noun',
  'DifVal'            : 0
},
{
  'Jenis Pemeriksaan' : 'LDL Cholesterol',
  'Hasil Pemeriksaan' : 180,
  'Interpretasi'      : 'Meningkat',
  'I_type'            : 'Verb',
  'DifVal'            : 51
},
{
  'Jenis Pemeriksaan' : 'Trygliserida',
  'Hasil Pemeriksaan' : 200,
  'Interpretasi'      : 'Normal',
  'I_type'            : 'Noun',
  'DifVal'            : 0
},
{
  'Jenis Pemeriksaan' : 'HBsAG',
  'Hasil Pemeriksaan' : 'Non Reaktif',
  'Interpretasi'      : 'Normal',
  'I_type'            : 'Noun',
  'DifVal'            : 0
},
}
```

**Gambar 5.** Keluaran proses interpretasi data (*Interpreted Data*)

**Tabel 1.** Rentang nilai normal kimia darah [9]

Pemeriksaan	Satuan	Rentang Nilai	
		Pria	Wanita
Glukosa Puasa	mg/dl	70 - 110	
SGOT	μ/L	5 - 35	5 - 31
SGPT	μ/L	5 - 45	5 - 34
Ureum	mg/dl	10 - 50	
Creatinine	mg/dl	0,6 – 1,4	0,5 – 1,1
Asam Urat	mg/dl	3,5 – 7,2	2,9 – 5,2
Cholesterol Total	mg/dl	< 200	
HDL	mg/dl	35 - 55	35 - 70
LDL	mg/dl	< 130	
Trygliserida	mg/dl	< 150	

**Tabel 2.** Interpretasi komponen pemeriksaan kimia darah [9]

No	Interpretasi	Aturan
1	Normal (noun)	Jika nilai komponen hasil pemeriksaan kimia darah berada di antara rentang nilai pada <a href="#">Tabel 1</a> .
2	Meningkat / Peningkatan (verb / noun)	Jika nilai komponen hasil pemeriksaan hematologi berada di atas rentang batas tertinggi sesuai dengan data pada <a href="#">Tabel 1</a> .
3	Menurun / Penurunan (verb / noun)	Jika nilai komponen hasil pemeriksaan hematologi berada di bawah rentang batas terendah sesuai dengan data pada <a href="#">Tabel 1</a> .



*Determination* menggunakan metode seperti yang dinyatakan pada [Algoritme 3](#).

## E. Pembangkitan teks

Hasil keluaran *Data Determination* yang berupa *Selected Data* diteruskan ke pembangkitan teks (*Text Generation*) untuk membangkitkan kalimat interpretasi yang sesuai dengan *Selected Data*. Proses ini merupakan bagian penting pada penelitian ini karena bertanggung jawab dalam membangkitkan kalimat berdasarkan data terkait. Untuk membangkitkan kalimat, penelitian ini menggunakan pendekatan berbasis *r-template* (*rule-template based*). Pendekatan ini dipilih karena pada domain kasus ini, kombinasi kalimat yang dihasilkan tidak banyak sehingga pendekatan ini dipandang lebih efisien dan efektif dibandingkan menggunakan pendekatan probabilistik, seperti *n-gram*.

Namun, pendekatan *template* pada umumnya memberikan kesan kaku pada kalimat yang dihasilkan, terutama dalam mengabungkan beberapa kalimat untuk menjadi suatu paragraf. Selain itu, apabila di dalam *template* tersedia hanya untuk jumlah parameter tertentu sementara jumlah parameter dapat berubah-ubah jumlahnya, maka *template* yang dibangun haruslah banyak dengan struktur yang memiliki kemiripan besar. Untuk mengatasi masalah tersebut, penelitian ini menggunakan *r-template* dimana *template* yang dibangun tidak perlu banyak.

### Template corpora

*Template Corpora* merupakan kumpulan *template* yang terbagi dalam beberapa corpus. Corpus yang pertama berisi kalimat-kalimat yang dapat digunakan sebagai *template* untuk membangkitkan kalimat pembuka pada sistem ini. Sementara itu, corpus kedua berisi kalimat-kalimat yang lebih spesifik untuk menjelaskan kondisi yang lebih rinci terhadap hasil pemeriksaan yang telah dilakukan.

Pada penelitian ini, corpora yang dibangun masih dilakukan secara manual (*hand-crafted*) dengan melibatkan pakar ilmu patologi klinis. Pakar diminta menuliskan interpretasi singkat atas hasil pemeriksaan tertentu. Kalimat interpretasi tersebut diberikan penanda (*tag*) sebagai identitas suatu kata atau frase tersebut. Corpus pertama (Corpus1) disebut corpus kalimat pembuka. Sebelum menjelaskan interpretasi secara lebih spesifik, penjelasan tersebut diawali dengan interpretasi umum yang direpresentasikan sebagai kalimat pembuka penjelasan hasil pemeriksaan kimia darah.

Dalam kasus ini, kalimat pembuka hanya terdiri atas 4 (empat) kombinasi kalimat atau hanya terdapat 4 (empat) kandidat *template* untuk kalimat pembuka. Corpus ini diberikan 3 (tiga) identitas, yakni NAMA PASIEN, JUM\_ABPAR, dan ABPAR. Keempat *template* tersebut adalah sebagai berikut:

- Berdasarkan hasil pemeriksaan laboratorium diketahui bahwa terdapat

```
[
{
  'Nama Pasien' : 'Akbar Duwi',
  'Jenis Kelamin': 'Pria',
  'Jum_Abpar' : 2
  'Pemeriksaan' :
    [
      {
        'Jenis Pemeriksaan' : 'Ureum',
        'Hasil Pemeriksaan' : 90,
        'Interpretasi' : 'Meningkat',
        'I_type' : 'Verb',
        'DifVal' : 40
      },
      {
        'Jenis Pemeriksaan' : 'LDL Cholesterol',
        'Hasil Pemeriksaan' : 180,
        'Interpretasi' : 'Meningkat',
        'I_type' : 'Verb',
        'DifVal' : 51
      }
    ]
  }
]
```

Gambar 6. Keluaran proses penentuan data (*Selected Data*)

### Algoritme 2. Pseudocode interpretasi data

```
1 Input : Extracted Data as ED
2 Output: Interpreted Data
3 Function DataInterpretation(ED)
4   foreach pemeriksaan(value) ∈ ED(pemeriksaan)
5     if Jenis Kelamin = Pria
6       if Hasil Pemeriksaan(Jenis Pemeriksaan) <
         Lower Bound(Pria) Then interpretasi
         = 'penurunan' or 'menurun' and
         I_type = 'noun' or 'verb'
7       if Hasil Pemeriksaan(Jenis Pemeriksaan) >
         Upper Bound(Pria) Then interpretasi
         = 'peningkatan' or 'meningkat'
         and I_type = 'noun' or 'verb'
8       else
9         interpretasi = 'normal' and
          I_type = 'noun'
10      else
11        if Hasil Pemeriksaan(Jenis Pemeriksaan) <
          Lower Bound(wanita) Then interpretasi
          = 'penurunan' or 'menurun' and
          I_type = 'noun' or 'verb'
12        if Hasil Pemeriksaan(Jenis Pemeriksaan) >
          Upper Bound(wanita) Then interpretasi
          = 'peningkatan' or 'meningkat'
          and I_type = 'noun' or 'verb'
13        else
14          interpretasi = 'normal' and
           I_type = 'noun'
15      Hitung besaran selisih nilai as difVal
16    end foreach
17 end function
```

### Algoritme 3. Pseudocode penentuan data

```
1 Input : Interpreted Data as ID
2 Output: Selected Data
3 Function DataDetermination(ID)
4   foreach pemeriksaan(value) ∈ ID
5     if (DifVal for Jenis Pemeiksaan) = 0
6       Remove Jenis Pemeriksaan from
        pemeriksaan(value)
7     else
8       continue
9   end for
10 Count pemeriksaan(value) as Jum_AbPar
11 end function
```

- [NUM\_ABPAR] komponen pemeriksaan yang abnormal, yakni [ABPAR]
- Berdasarkan hasil pemeriksaan laboratorium yang telah dilakukan, diketahui ada [NUM\_ABPAR] komponen pemeriksaan yang abnormal, yakni [ABPAR]
- Hasil pemeriksaan laboratorium [NAMA PASIEN] menunjukkan ada [NUM\_ABPAR] komponen pemeriksaan yang abnormal, yakni [ABPAR]
- Berdasarkan hasil pemeriksaan laboratorium [NAMA PASIEN], terdapat [NUM\_ABPAR] komponen pemeriksaan yang abnormal, yaitu [ABPAR]

Corpus kedua (Corpus2) disebut dengan corpus kalimat rinci. Setelah disampaikan kalimat yang merupakan penjelasan secara umum terkait hasil pemeriksaan kimia darah, selanjutnya harus diikuti dengan kalimat yang menginterpretasikan secara rinci kondisi setiap parameter yang abnormal. Dalam kasus ini, penelitian ini menggunakan 9 (sembilan) kombinasi kalimat, sehingga menghasilkan 9 (sembilan) kandidat *template* untuk menjelaskan kondisi rinci dari komponen hasil pemeriksaan darah yang abnormal. Corpus ini dibangun dengan memanfaatkan beberapa identitas seperti berikut:

- Nilai komponen pemeriksaan [JENIS PEMERIKSAAN] mengalami [I\_NOUN] sebesar [DIFVAL] poin dari batas normal
- Nilai [JENIS PEMERIKSAAN] mengalami [I\_NOUN] sebesar [DIFVAL] poin dari batas normal
- Nilai komponen pemeriksaan [JENIS PEMERIKSAAN] mengalami [I\_NOUN] sebesar [DIFVAL] poin
- Nilai [JENIS PEMERIKSAAN] mengalami [I\_NOUN] sebesar [DIFVAL] poin
- [I\_NOUN] sebesar [DIFVAL] poin dari batas normal terjadi pada nilai komponen pemeriksaan [JENIS PEMERIKSAAN]
- Nilai komponen pemeriksaan [JENIS PEMERIKSAAN] [I\_VERB] sebesar [DIFVAL] poin dari batas normal
- Nilai [JENIS PEMERIKSAAN] [I\_VERB] sebesar [DIFVAL] poin dari batas normal
- Nilai komponen pemeriksaan [JENIS PEMERIKSAAN] [I\_VERB] sebesar [DIFVAL] poin
- Nilai [JENIS PEMERIKSAAN] [I\_VERB] sebesar [NILAI\_PERUBAHAN] poin

#### Template determination

Pembangkitan *template* (*Template Determination*) merupakan tahap pertama dari *Text Generation*. Tahap ini bertanggung jawab untuk melakukan pemilihan kandidat *template* dari *template corpora* yang sesuai dengan konten informasi yang akan disampaikan pada laporan tekstual. Masukan ke proses ini adalah *Selected Data* yang merupakan keluaran dari proses *Data Determination*. Proses pemilihan dan penentuan *template* yang digunakan pada penelitian ini adalah dengan menggunakan pendekatan acak (*random-based*), sebagaimana yang disajikan pada [Algoritme 4](#). Contoh hasil dari proses *Template Determination* yang mengacu pada *Selected Data* di [Gambar 6](#), dimana Num\_AbPar

- ListTemplate =
- [
- "Berdasarkan hasil pemeriksaan laboratorium [NAMA PASIEN], terdapat [JUM\_ABPAR] komponen pemeriksaan yang abnormal, yaitu [ABPAR]",
- "Nilai komponen pemeriksaan [PAR] [I\_VERB] sebesar [NILAI\_PERUBAHAN] poin dari batas normal",
- "Nilai [PAR] [I\_VERB] sebesar [NILAI\_PERUBAHAN] poin dari batas normal"
- ]

Gambar 7. Keluaran proses penentuan template

#### Algoritme 4. Pseudokode penentuan template

```

1 Input : Selected Data as SD
2 Output: List of Template
3 Function TemplateDetermination(SD(Num_AbPar))
4   ListTemplate = []
5   Rand Template ∈ Corpus1
6   Insert SelectedTemplate into ListTemplate
7   for i=0 to Num_AbPar-1 do
8     Rand Template ∈ Corpus2 based on
       SD(Pemeriksaan(i.type))
9     Insert SelectedTemplate into
       ListTemplate
10  end for
11 end function

```

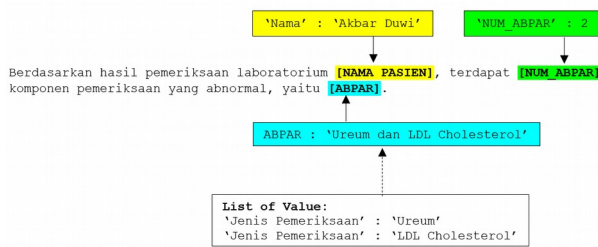
teridentifikasi sebanyak 2 (dua) buah, seperti ditunjukkan pada [Gambar 7](#).

#### Data assignment

Setelah memperoleh ListTemplate yang sesuai, tahap selanjutnya adalah penugasan data (*Data Assignment*). Tahap ini merupakan tahap dimana setiap slot pada *template* yang ada dalam ListTemplate akan digantikan dengan data yang sesuai pada *Selected Data*. Namun, proses penggantian slot untuk kalimat umum menerapkan pendekatan berbasis aturan, terutama untuk slot [ABPAR]. Aturan tersebut dimaksudkan untuk dapat sedikit meningkatkan fleksibilitas pembangkitan kalimat dengan menyesuaikan jumlah parameter abnormal. Konsep proses penggantian slot untuk kalimat umum tersebut dapat ditunjukkan seperti pada [Gambar 8](#).

Slot [ABPAR] menggunakan aturan konjungsi koordinatif berjenis penambahan melalui prosedur sebagaimana yang disajikan pada [Algoritme 5](#). Aturan ini akan bekerja pada sistem jika terdapat lebih dari 1 (satu) parameter yang abnormal. Strategi ini dimaksudkan agar jumlah *template* kalimat umum yang sama, namun berbeda jumlah slot parameter, dapat diminimalisir di dalam corpus. Dengan kondisi seperti itu, fleksibilitas jumlah parameter dapat tercapai dengan baik tanpa memberikan kesan kaku dalam ketersediaan *template*. Keluaran dari tahap ini adalah suatu representasi tekstual untuk setiap konten atau kondisi yang akan disampaikan. Dengan menggunakan metode

Selected Template:



**Gambar 8.** Mekanisme pembangkitan kalimat umum berbasis aturan untuk slot ABPAR

#### Algoritme 5. Pseudocode [ABPAR] aggregation

```

1  Input  : Template [0] and Selected Data
2  Output: Aggregated Sentences
3  Function AbPar_Aggregation(AbParName[])
4      If len(AbParName[]) == 2 Then
5          AbPar = Aggregate AbParName[0] dan
                    AbParName[1] dengan kata "dan"
6      Else
7          W1 = Aggregate AbParName[0] to
                    AbParName[len(AbParName[])-1]
                    dengan simbol ", "
8          AbPar = Aggregate W1 dan
                    AbParName[len(AbParName[])]
                    dengan kata "dan"
9      return AbPar
10 end function
    
```

yang ada pada [Algoritme 1](#) dan [Algoritme 2](#), hasil proses *Data Assignment* berdasarkan *Selected Data* di [Gambar 6](#) dan *Selected Template* di [Gambar 7](#) dapat ditunjukkan seperti pada [Gambar 9](#).

#### Sentence aggregation

Representasi tekstual awal yang dihasilkan oleh tahap *Data Assignment* diteruskan ke tahap agregasi kalimat (*Sentence Aggregation*). Tahap ini bertujuan untuk menggabungkan dua atau lebih kalimat untuk menjadi suatu paragraf termasuk dengan menambahkan tanda baca yang relevan. Keluaran proses ini dapat dihasilkan dengan menggunakan konsep dasar seperti ditunjukkan pada [Gambar 10](#).

Proses *Sentence Aggregation* pada penelitian ini menggunakan metode penggabungan sederhana dua atau lebih kalimat sebagaimana yang disajikan pada [Algoritme 6](#). Hasil dari proses ini digunakan sebagai masukan pada tahap *Report Generation*.

#### Report generation

Proses terakhir dari *Text Generation* adalah *Report Generation*. Proses ini dimaksudkan untuk menyisipkan *Generated Sentences* ke dalam dokumen hasil pemeriksaan darah. Hasil keluaran dari proses ini adalah suatu dokumen hasil pemeriksaan darah yang telah dilengkapi dengan penjelasan tekstual yang dibangkitkan. Keluaran tersebut dapat dilihat seperti pada [Gambar 11](#).

1. ListTemplate =
2. [
3. "Berdasarkan hasil pemeriksaan laboratorium Akbar Duwi, terdapat 2 komponen pemeriksaan yang abnormal, yaitu Ureum dan LDL Cholesterol",
4. "Nilai komponen pemeriksaan Ureum meningkat sebesar 40 poin dari batas normal",
5. "Nilai LDL Cholesterol meningkat sebesar 51 poin dari batas normal"
6. ]

**Gambar 9.** Keluaran proses penugasan data

1. GeneratedSentences =
2. "Berdasarkan hasil pemeriksaan laboratorium Akbar Duwi, terdapat 2 komponen pemeriksaan yang abnormal, yaitu Ureum dan LDL Cholesterol. Nilai komponen pemeriksaan Ureum meningkat sebesar 40 poin dari batas normal. Nilai LDL Cholesterol meningkat sebesar 51 poin dari batas normal."

**Gambar 10.** Keluaran proses agregasi kalimat

#### Algoritme 6. Pseudocode agregasi kalimat

```

1  Input  : List Template Data Assignment as LTDA
2  Output: Generated Sentences
3  Function SentenceAggregation(LTDA)
4      GeneratedSentence = " "
5      for i=0 to length(LTDA) - 1 do
6          GeneratedSentences = LTDA[i] + " ."
7      end for
8  end function
    
```

Nama Pasien : Akbar Duwi  
Tanggal Lahir : 4 Maret 1979  
Jenis Kelamin : Pria

Dokter Pengirim : dr. Dina Mutia  
Tanggal Pemeriksaan : 26 Juli 2016  
Nomor Pemeriksaan : KD260716022

JENIS PEMERIKSAAN	HASIL PEMERIKSAAN	NILAI RUJUKAN
<b>KIMIA DARAH</b>		
<b>DIABETES</b>		
Glucose	86	70 - 180 mg/dl
<b>FUNGSI HATI</b>		
SGOT	24	W : ≤ 31 u/l P : ≤ 35 u/l
SGPT	26	W : ≤ 34 u/l P : ≤ 45 u/l
<b>FUNGSI JENJAL</b>		
Ureum	90	10 - 50 mg/dl
Creatinine	1.1	W : 0.5 - 1.1 mg/dl P : 0.6 - 1.4 mg/dl
Asam Urat	5.2	P : 3.5 - 7.2 mg/dl W : 2.9 - 5.2 mg/dl
<b>LEMAK</b>		
Cholesterol Total	290	< 200 mg/dl
HDL Cholesterol	39	W : 35 - 70 mg/dl P : 35 - 55 mg/dl
LDL Cholesterol	180	< 130 mg/dl
Trigliserida	200	< 150 mg/dl
<b>IMUNOLOGI</b>		
HBsAg	NON REAKTIF	Negatif

Catatan :  
Berdasarkan hasil pemeriksaan laboratorium Akbar Duwi, terdapat 2 komponen pemeriksaan yang abnormal, yaitu Ureum dan LDL Cholesterol. Nilai komponen pemeriksaan Ureum meningkat sebesar 40 poin dari batas normal. Nilai LDL Cholesterol meningkat sebesar 51 poin dari batas normal.

Pemeriksa,

dr. Muhammad Putra, Sp.PK

**Gambar 11.** Keluaran BTISys

### III. HASIL DAN PEMBAHASAN

Tujuan dari penelitian ini adalah membangun suatu sistem yang mampu menginterpretasikan data hasil pemeriksaan kimia darah ke dalam suatu representasi tekstual berbahasa Indonesia. Penelitian ini menggunakan 20 dokumen berekstensi Microsoft Word yang berisi data hasil pemeriksaan kimia darah. Dokumen tersebut diambil dari salah satu rumah sakit di Kota Medan.

#### A. Penilaian

Keberhasilan dan keandalan BTISys dinilai berdasarkan tingkat *naturalness* dari kalimat yang dibangkitkan oleh BTISys. Tingkat *naturalness* ini dievaluasi dengan mempertimbangkan 3 (tiga) aspek, yaitu aspek *readability*, *clarity*, dan *general appropriateness* seperti [7], [8], [10]. Penilaian ini dilakukan oleh dokter muda atau dokter koas yang merupakan pengguna utama dari keluaran BTISys melalui suatu kuisioner terstruktur.

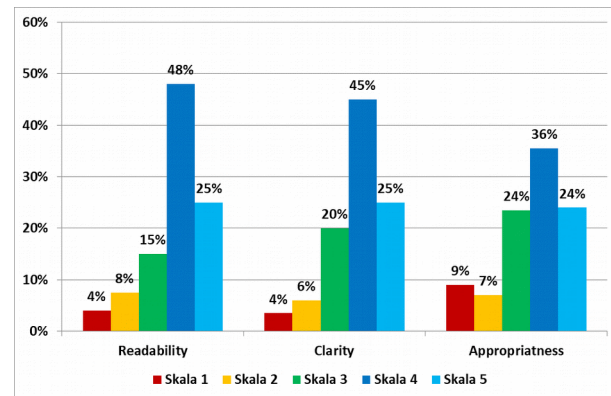
Aspek *readability* merupakan aspek yang digunakan untuk mengetahui tingkat kemudahan pengguna dalam membaca kalimat yang dibangkitkan oleh BTISys sehingga berdampak pada tingkat pemahaman pengguna secara cepat dan tepat. Penilaian terhadap aspek ini terbagi atas 5 (lima) skala, yaitu:

- Skala 1: kalimat yang dibangkitkan sangat sulit dibaca sehingga pengguna sangat sulit untuk memahaminya secara cepat dan tepat.
- Skala 2: kalimat yang dibangkitkan sulit dibaca sehingga pengguna sulit untuk memahaminya secara cepat dan tepat.
- Skala 3: kalimat yang dibangkitkan cukup mudah dibaca sehingga pengguna cukup mudah untuk memahaminya secara cepat dan tepat.
- Skala 4: kalimat yang dibangkitkan mudah dibaca sehingga pengguna mudah memahaminya secara cepat dan tepat.
- Skala 5: kalimat yang dibangkitkan sangat mudah dibaca sehingga pengguna sangat mudah untuk memahaminya secara cepat dan tepat.

Aspek kedua yang dinilai adalah aspek *clarity*. Aspek ini digunakan untuk mengukur tingkat kejelasan penyampaian informasi ke dalam suatu representasi kalimat dalam bahasa Indonesia yang mampu ditangkap oleh pengguna BTISys. Penilaian terhadap aspek ini terdiri atas 5 (lima) skala, yaitu:

- Skala 1: kalimat yang dibangkitkan memberikan informasinya yang sangat tidak jelas.
- Skala 2: kalimat yang dibangkitkan memberikan informasinya yang tidak jelas.
- Skala 3: kalimat yang dibangkitkan memberikan informasinya yang cukup jelas.
- Skala 4: kalimat yang dibangkitkan memberikan informasinya yang jelas.
- Skala 5: kalimat yang dibangkitkan memberikan informasinya yang sangat jelas.

Aspek terakhir yang dinilai adalah aspek *general appropriateness*. Aspek ini merupakan aspek yang



Gambar 12. Hasil penilaian dokter muda atau koas

digunakan untuk menilai apakah kalimat yang dibangkitkan oleh BTISys telah memiliki kesesuaian yang baik sehingga bisa mencukupi atau menambah pengetahuan pengguna secara cepat dan tepat. Penilaian ini sama halnya dengan penilaian lainnya yang terbagi atas 5 (lima) skala, yaitu:

- Skala 1: kalimat yang dibangkitkan sangat tidak sesuai atau tepat.
- Skala 2: kalimat yang dibangkitkan tidak sesuai atau tepat.
- Skala 3: kalimat yang dibangkitkan cukup sesuai atau tepat.
- Skala 4: kalimat yang dibangkitkan sesuai atau tepat.
- Skala 5: kalimat yang dibangkitkan sangat sesuai atau tepat.

#### B. Hasil pengujian

Kuisioner diberikan kepada 10 orang dokter muda atau koas yang merupakan pengguna utama dari sistem BTISys. Masing-masing dokter tersebut diberikan sebuah paket kuisioner yang berisi 20 hasil pemeriksaan kimia darah yang masing-masing hasil tersebut telah dilengkapi dengan interpretasi tekstual berbahasa Indonesia yang dibangkitkan oleh BTISys. Dengan kata lain, setiap 1 orang dokter muda atau koas menilai 20 interpretasi tekstual yang dihasilkan oleh BTISys untuk satu paket kuesioner.

Dokter muda tersebut diminta untuk membaca interpretasi tekstual dimaksud berdasarkan 3 (tiga) aspek penilaian. Melalui pengujian ini, hasil dari penilaian yang diberikan oleh 10 dokter muda atau koas terhadap hasil pembangkitan kalimat dari BTISys disajikan pada Gambar 12.

#### C. Pembahasan

Hasil penilaian yang dilakukan oleh dokter muda atau koas menunjukkan bahwa sistem BTISys andal dalam menghasilkan interpretasi tekstual berdasarkan data hasil pemeriksaan kimia darah. Hal ini terlihat dari penilaian dokter muda atau koas yang menyatakan bahwa tingkat *naturalness* kalimat interpretasi yang dibangkitkan oleh BTISys adalah 85,3 %. Secara



spesifik dokter muda atau koas lebih dominan menilai kalimat interpretasi yang dibangkitkan oleh BTISys dengan skala 4 dan 5 yang menyatakan bahwa kalimat yang dibangkitkan sudah disampaikan secara lugas dan tepat.

#### Readability

Berdasarkan hasil penilaian yang disajikan pada [Gambar 12](#), dokter muda atau koas menilai bahwa kalimat yang dibangkitkan oleh BTISys dapat dibaca dengan mudah sehingga mudah untuk dipahami. Hal ini disebabkan karena kalimat yang dibangkitkan oleh BTISys terdiri atas kumpulan kata-kata yang disusun secara koheren dan memiliki hubungan antar kata yang saling berkaitan sesuai [11]. Susunan ini disebabkan aturan penyusunan antar kata di BTISys telah diformulasikan mengikuti aturan tata bahasa Indonesia yang mengikuti [12].

Keberhasilan sistem BTISys lebih bergantung kepada corpus yang dibangun pada penelitian ini karena pembangkitan kalimatnya berbasis korpus yang sesuai [13]. Dengan kata lain, BTISys tidak dapat membangkitkan kalimat tanpa adanya dukungan korpus yang berisi contoh kalimat di suatu domain tertentu. Oleh karena itu, korpus yang dirancang telah mengikuti kaidah tata bahasa Indonesia pada umumnya dengan sedikit melanggar kaidah tata bahasa agar dapat memberikan capaian *naturalness* yang baik.

Dengan korpus tersebut, kalimat yang dibangkitkan dapat tersusun secara *well-formedness* atau mendekati *well-formedness* sehingga bisa memberikan dampak terhadap tingkat *naturalness* suatu bahasa sesuai [12], [14], terutama pada aspek *readability* pembaca. Meskipun beberapa dokter muda atau koas masih menilai 15 % kalimat yang dibangkitkan cukup mudah dibaca dan 12 % tidak dapat dibaca, persentase tersebut tidak serta merta menyatakan bahwa kalimat yang dibangkitkan memiliki kedudukan yang sulit dibaca karena persentasenya jauh di bawah dari 73 %.

#### Clarity

Aspek *clarity* menekankan kepada kejelasan penyampaian informasi yang disajikan dalam suatu representasi kalimat. Penilaian dokter muda atau koas terhadap aspek ini menunjukkan bahwa BTISys sudah andal dalam membangkitkan kalimat yang memiliki kejelasan yang baik. Hal ini disebabkan aspek ini dapat dicapai apabila aspek *readability* terpenuhi. Kalimat yang tersusun dengan baik memberikan pemahaman yang baik pula kepada pembaca sehingga informasi yang disampaikan dalam kalimat tersebut dapat secara jelas ditangkap oleh pembaca. Oleh karena itu, persentase aspek *readability* untuk skala 4 dan 5 tidak jauh berbeda dengan aspek *clarity* yang hanya memiliki selisih 3 %.

#### General appropriateness

Aspek ini menekankan kepada penilaian apakah kalimat interpretasi yang dibangkitkan oleh BTISys

telah mencukupi kesesuaian informasi sehingga berguna menambahkan pengetahuan pengguna. Dokter muda atau koas telah menilai bahwa kalimat interpretasi yang dibangkitkan telah sesuai atau tepat untuk disampaikan. Hal ini terlihat bahwa dokter muda atau koas menilai 60 % kalimat interpretasi tersebut telah sesuai atau tepat.

Meskipun demikian, sistem ini masih dipandang perlu adanya pemuktahiran agar aspek ini dapat dipenuhi secara lebih baik lagi. Kondisi ini disebabkan masih adanya pengguna yang menilai tingkat kesesuaian kalimat interpretasi yang dibangkitkan sebesar 24 % dan 16 %. Penilaian ini diberikan oleh dokter muda atau koas karena kalimat yang dibangkitkan sudah sesuai kebutuhan yang diharapkan berdasarkan kondisi interpretasi yang tepat sebagaimana yang dijelaskan dalam [9], [15]. Namun, dokter muda atau koas masih berharap adanya informasi tentang kondisi urgensi yang lebih rinci dalam interpretasi hasil pemeriksaan kimia darah sehingga nantinya dapat juga membantu meningkatkan pengetahuan terkait kimia darah secara lebih mendalam.

Penelitian selanjutnya dapat dilakukan untuk mengembangkan BTISys dengan meningkatkan fleksibilitas *template* dan pengetahuan para pakar, baik dalam hal ilmu medis maupun ilmu membangkitkan kalimat. Selain itu, pendekatan yang lebih baik dapat dikembangkan dengan evaluasi tidak terbatas pada penilaian manusia, namun dilakukan secara otomatis.

## IV. KESIMPULAN

Sistem yang diusulkan dapat digunakan untuk membangkitkan interpretasi data hasil pemeriksaan kimia darah dalam suatu representasi tekstual berbahasa Indonesia. Penilaian pengguna yang memandang kalimat yang dibangkitkan dengan menggunakan pendekatan berbasis *corpus* dengan *r-template based* memiliki tingkat *naturalness* yang baik.

## DAFTAR PUSTAKA

- [1] National Heart, Lung, and Blood Institute, "Blood tests," NHBLI, 2012. [Online]. Available: <https://www.nhlbi.nih.gov/health/health-topics/topics/bdt>. [Accessed: 2 Dec. 2018].
- [2] I. Adeyanju, "Generating weather forecast texts with case based reasoning," *International Journal of Computer Applications*, vol. 45, no. 10, pp. 35-40, 2012. doi: [10.5120/6819-9176](https://doi.org/10.5120/6819-9176)
- [3] M. Molina, J. Sanchez-Soriano, and O. Corcho, "Using open geographic data to generate natural language descriptions for hydrological sensor networks," *Sensors*, vol. 15, no. 7, pp. 16009-16026, 2015. doi: [10.3390/s150716009](https://doi.org/10.3390/s150716009)
- [4] J. Mahapatra, S. K. Naskar, and S. Bandyopadhyay, "Statistical natural language generation from tabular non-textual data," in *9th International Natural Language Generation Conference*, Edinburg, UK, Sept. 2016, pp. 143-153.

- [5] A. Gatt et al., "From data to text in the neonatal intensive care unit: Using NLG technology for decision support and information management," *AI Communications*, vol. 22, no. 3, pp. 153-186, 2009. doi: [10.3233/AIC-2009-0453](https://doi.org/10.3233/AIC-2009-0453)
- [6] B. Di Eugenio et al., "PatientNarr: towards generating patient-centric summaries of hospital stays," in *8th International Natural Language Generation Conference*, Philadelphia, USA, Jun. 2014, pp. 6-10. doi: [10.3115/v1/W14-4402](https://doi.org/10.3115/v1/W14-4402)
- [7] W. Pratomo and A. M. Barmawi, "Performing chart interpretation using lexical selection in Indonesian language," in *2013 International Conference on Advanced Computer Science and Information Systems*, Bali, Indonesia, Sept. 2013, pp. 199-205. doi: [10.1109/ICACSYS.2013.6761576](https://doi.org/10.1109/ICACSYS.2013.6761576)
- [8] I. Aulia and A. M. Barmawi, "An automatic health surveillance chart interpretation system based on Indonesian language," in *2015 International Conference on Advanced Computer Science and Information Systems*, Depok, Indonesia, Oct. 2015, pp. 163-170. doi: [10.1109/ICACSYS.2015.7415165](https://doi.org/10.1109/ICACSYS.2015.7415165)
- [9] Kementerian Kesehatan Republik Indonesia, *Pedoman interpretasi data klinik*. Jakarta: Kementerian Kesehatan Republik Indonesia, 2011.
- [10] A. Belz and E. Reiter, "Comparing automatic and human evaluation of NLG systems," in *11th Conference of the European Chapter of the Association for Computational Linguistics*, Trento, Italy, Apr. 2006, pp. 213-320.
- [11] R. Vadlapudi and R. Katragadda, "On automated evaluation of readability of summaries: capturing grammaticality, focus, structure and coherence," in *NAACL HLT 2010 Student Research Workshop*, Los Angeles, USA, Jun. 2010, pp. 7-12.
- [12] K. Willems and L. D. Cuypere, *Naturalness and iconicity in language*. John Benjamins Publishing, 2008.
- [13] E. Reiter and S. Sripada, "Should corpora texts be gold standards for NLG?," in *International Natural Language Generation Conference*, New York, USA, Jul. 2002, pp. 97-104.
- [14] J. Sinclair, "Naturalness in language," *Ilha do Desterro*, vol. 5, no. 11, pp. 203-210, 1984.
- [15] Boston University School of Public Health, "Atherosclerosis," 2016. [Online]. Available: [http://sphweb.bumc.bu.edu/otlt/MPH-Modules/PH/PH709\\_Heart/](http://sphweb.bumc.bu.edu/otlt/MPH-Modules/PH/PH709_Heart/). [Accessed: 4 Jan. 2019].